EECS 245, Winter 2026

LEC 24 Singular Value Decomposition,
Towards PCA

→ Read: Ch. 10

Agenda

- Recap: singular value decomposition
- Low-rank approximation
  → e.g., image compression
- Principal components analysis
  → Dimensionality reduction

Announcements

- HW 11 out, due Tuesday
- Don't skip HW 11 or lab this week!
- No mock exam: instead, will post "Post-MT2 Practice Worksheet"
- If 80% of the class fills out **both** the official evals <u>and</u> the End-of-Semester survey, everyone gets 1% EC to their overall grade

$$X = \underbrace{\begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{3\sqrt{2}} & -\frac{1}{\sqrt{3}} & -\frac{2}{3} \\ \frac{1}{\sqrt{6}} & -\frac{1}{3\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{2}{3} \\ 0 & \frac{2\sqrt{2}}{3} & 0 & \frac{1}{3} \\ \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} & 0 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 15 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{bmatrix}}_{V^T}$$

X is a 4×3 → every X has an SVD

$$X = U_{n \times n} \Sigma_{n \times d} V^T_{d \times d}$$

columns of U are eigenvectors of XX^T

$$\begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \end{bmatrix}$$

$\sigma_i = \sqrt{\lambda_i}$ of X^TX (or XX^T)

columns of V are eigenvectors of X^TX

defining property:

$$X \vec{v}_i = \sigma_i \vec{u}_i$$

left singular vector

right singular vector

$$U = \begin{bmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \\ | & | & & | \end{bmatrix} \qquad V = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_d \\ | & & | \end{bmatrix}$$

Example: Find singular values of

$$X = \begin{bmatrix} 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$$

Hint: X^TX and XX^T have the same λ_i's — be efficient!

$$XX^T = \begin{bmatrix} 10 & 12 \\ 12 & 17 \end{bmatrix}$$

σ_i's of X are $\sqrt{\lambda_i}$'s of XX^T

eigenvalues of XX^T : 26 and 1

so $\sigma_1 = \sqrt{26}$, $\sigma_2 = \sqrt{1} = 1$

sort σ_i's in decreasing order

$$X = \begin{bmatrix} 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}, \quad XX^T = \begin{bmatrix} 10 & 12 \\ 12 & 17 \end{bmatrix}$$

$$X = U\Sigma V^T$$

$$\Sigma = \begin{bmatrix} \sqrt{26} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$U = \begin{bmatrix} XX^T & XX^T \\ \text{eigvec} & \text{eigvec} \\ \lambda=26 & \lambda=1 \end{bmatrix} \qquad V = \begin{bmatrix} X^TX & X^TX & X^TX \\ \text{eigvec} & \text{eigvec} & \text{eigvec} \\ \lambda=26 & \lambda=1 & \lambda=0 \end{bmatrix}_{3\times 3}$$

U, V must be orthogonal matrices!

basis for $\operatorname{colsp}(X)$

basis for $\operatorname{nullsp}(X^T)$

number of non-zero σ_i's $= \operatorname{rank}(X) = r$

basis for $\operatorname{colsp}(X^T)$

basis for $\operatorname{nullsp}(X)$

$$X = \underbrace{\begin{bmatrix} | & \cdots & | & | & \cdots & | \\ \vec{u}_1 & \cdots & \vec{u}_r & \vec{u}_{r+1} & \cdots & \vec{u}_n \\ | & \cdots & | & | & \cdots & | \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sigma_1 & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \sigma_r & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} - & \vec{v}_1^T & - \\ & \vdots & \\ - & \vec{v}_r^T & - \\ - & \vec{v}_{r+1}^T & - \\ & \vdots & \\ - & \vec{v}_d^T & - \end{bmatrix}}_{V^T}$$

$\vec{u}_1, \ldots, \vec{u}_n$ eigenvectors of XX^T

$\vec{v}_1, \ldots, \vec{v}_d$ (columns of V) eigenvectors of X^TX

$$\underbrace{\begin{bmatrix} 3 & 2 & 5 \\ 2 & 3 & 5 \\ 2 & -2 & 0 \\ 5 & 5 & 10 \end{bmatrix}}_{X} = \underbrace{\begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{3\sqrt{2}} & -\frac{1}{\sqrt{3}} & -\frac{2}{3} \\ \frac{1}{\sqrt{6}} & -\frac{1}{3\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{2}{3} \\ 0 & \frac{2\sqrt{2}}{3} & 0 & \frac{1}{3} \\ \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} & 0 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 15 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{bmatrix}}_{V^T}$$

The summation view of the SVD says that:

$$X = \overset{\sigma_1}{15} \overset{\vec{u}_1}{\begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ 0 \\ \frac{2}{\sqrt{6}} \end{bmatrix}} \overset{\vec{v}_1^{\,T}}{\begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{bmatrix}} + \overset{\sigma_2}{3} \overset{\vec{u}_2}{\begin{bmatrix} \frac{1}{3\sqrt{2}} \\ -\frac{1}{3\sqrt{2}} \\ \frac{2\sqrt{2}}{3} \\ 0 \end{bmatrix}} \overset{\vec{v}_2^{\,T}}{\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}}$$

$$= \underbrace{\begin{bmatrix} \frac{5}{2} & \frac{5}{2} & 5 \\ \frac{5}{2} & \frac{5}{2} & 5 \\ 0 & 0 & 0 \\ 5 & 5 & 10 \end{bmatrix}}_{\text{rank-one matrix}} + \underbrace{\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 2 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\text{rank-one matrix}}$$

SVD says that

any X with rank r

can be written

$$X = \underbrace{\sigma_1 \vec{u}_1 \vec{v}_1^T}_{\text{outer product, rank one}} + \sigma_2 \vec{u}_2 \vec{v}_2^T + \cdots + \sigma_r \vec{u}_r \vec{v}_r^T$$

$$= \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T$$

low-rank approximation:

suppose $r = \operatorname{rank}(X)$,

$k \le r$

$$\underbrace{X_k}_{\text{rank-}k\text{ appnx}} = \sum_{i=1}^{k} \underbrace{\sigma_i \vec{u}_i \vec{v}_i^{T}}_{\text{sum of } k \text{ rank-one matices}}$$

Fact: X_k, as defined here, is the "best" rank-X approx of X, it minimizes

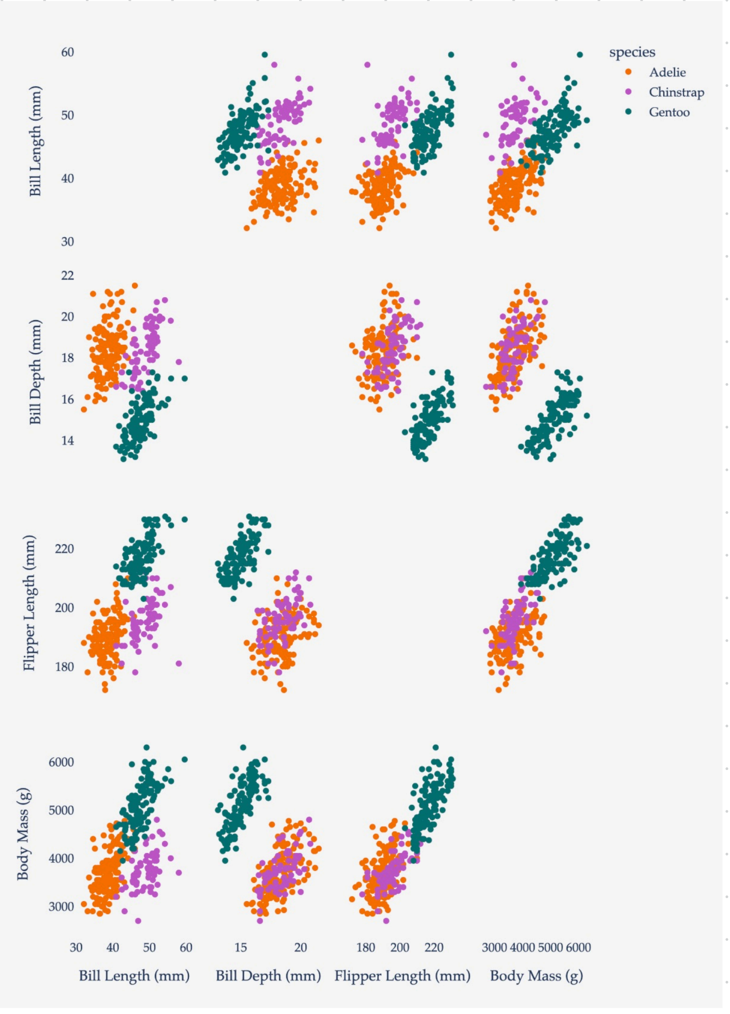
$$\|X - X_k\|_F$$

Frobenius norm!

Chapter 10.3

each row : 1 penguin

species	island	bill_length_mm	bill_depth_mm	flipper_length_mm	body_mass_g	sex
Adelie	Torgersen	39.1	18.7	181.0	3750.0	Male
Adelie	Torgersen	39.5	17.4	186.0	3800.0	Female
Adelie	Torgersen	40.3	18.0	195.0	3250.0	Female
Adelie	Torgersen	36.7	19.3	193.0	3450.0	Female
Adelie	Torgersen	39.3	20.6	190.0	3650.0	Male
...	...	...	...	...	...	...
Gentoo	Biscoe	47.2	13.7	214.0	4925.0	Female
Gentoo	Biscoe	46.8	14.3	215.0	4850.0	Female
Gentoo	Biscoe	50.4	15.7	222.0	5750.0	Male
Gentoo	Biscoe	45.2	14.8	212.0	5200.0	Female
Gentoo	Biscoe	49.9	16.1	213.0	5400.0	Male



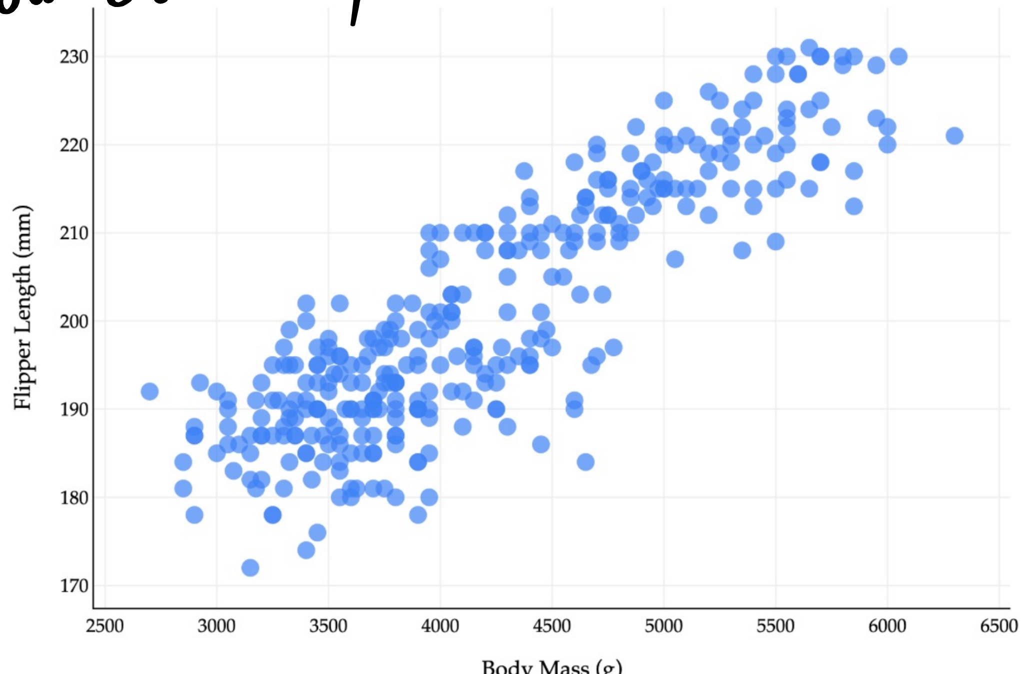
annoying to deal with 4 numerical features

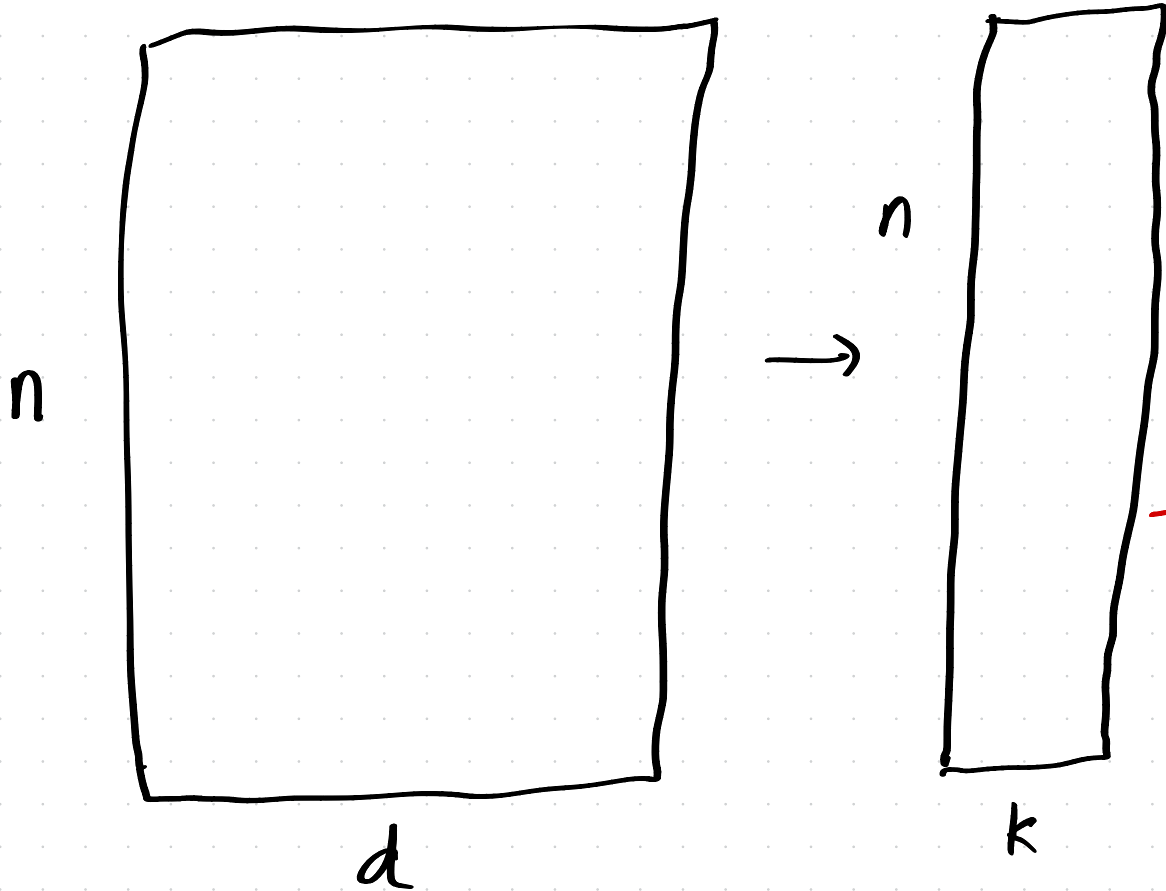
let's invent one good new feature

Goal: Turn data from $\mathbb{R}^2 \to \mathbb{R}^1$

$\text{new feature}_i = a\,(\text{body mass})_i + b\,(\text{flipper length})_i$

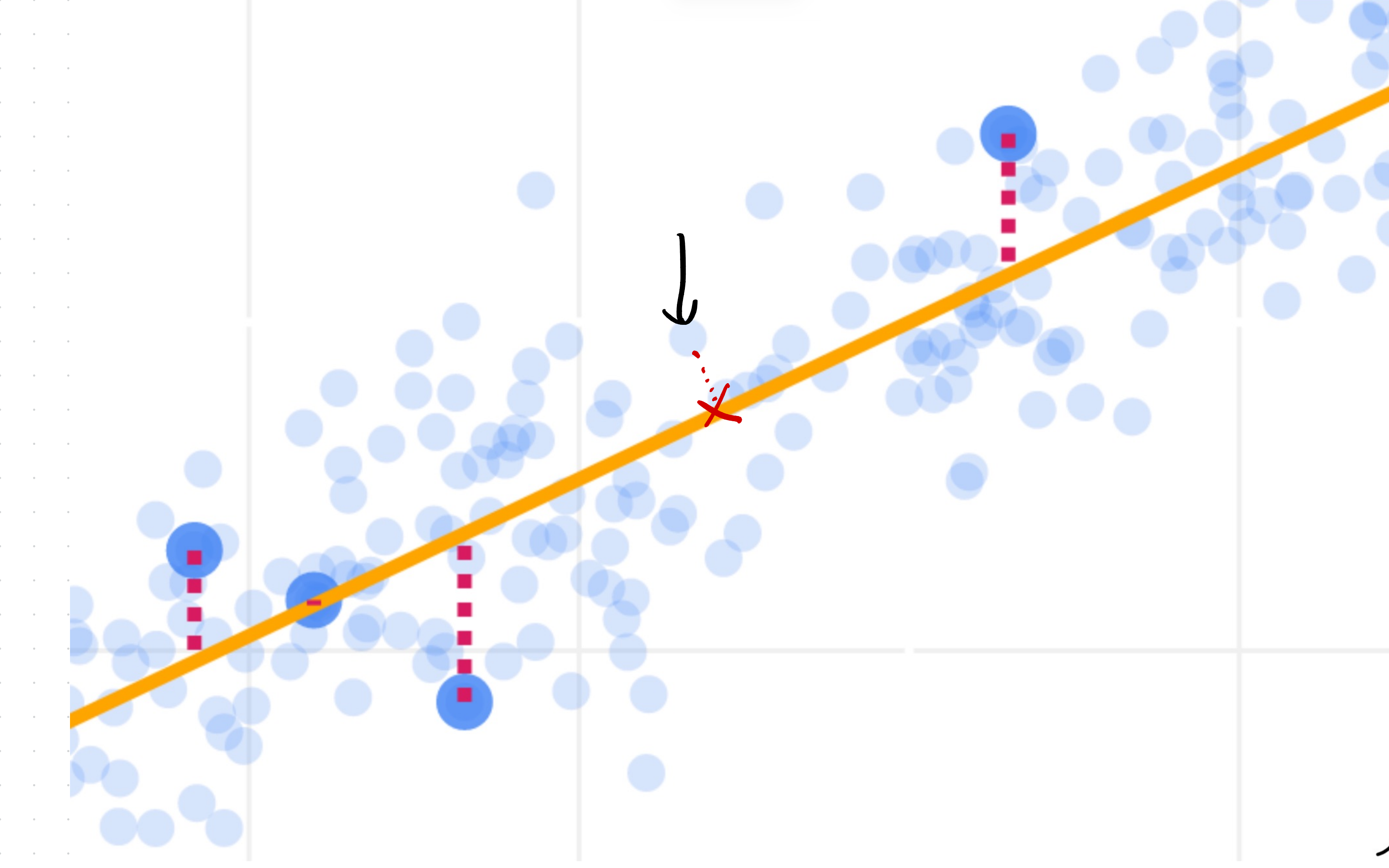
Q: How do we find a and b?



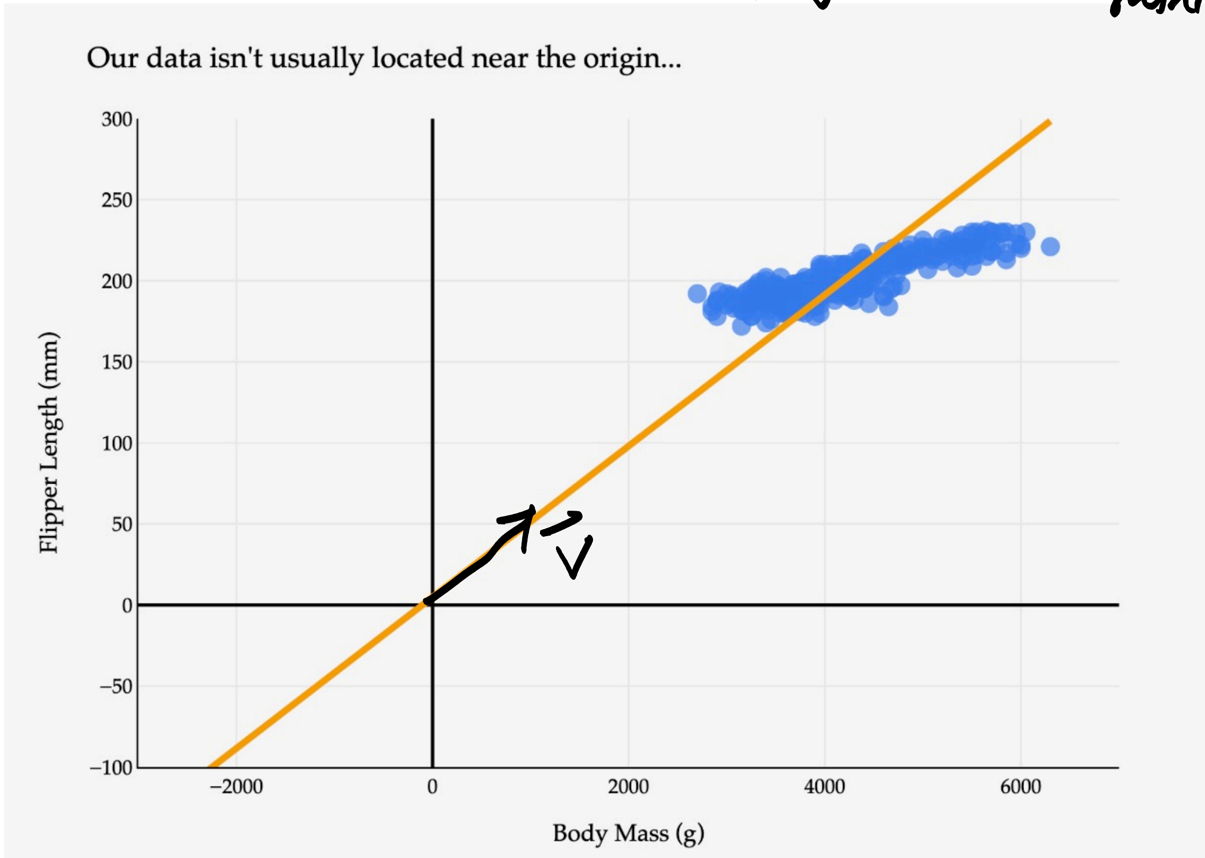


dimensionality reduction

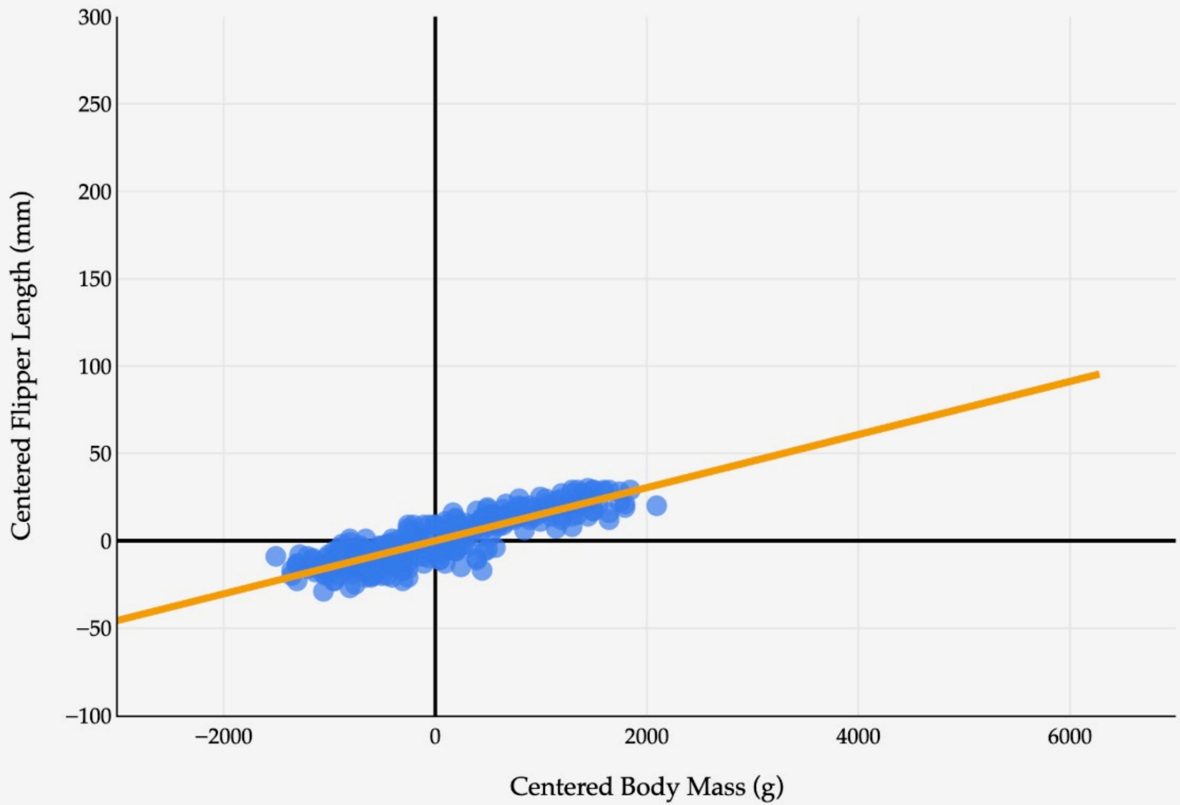
unsupervised learning



Goal: Find vector $\vec{v}$ to project each data point onto so that total projection error minimized

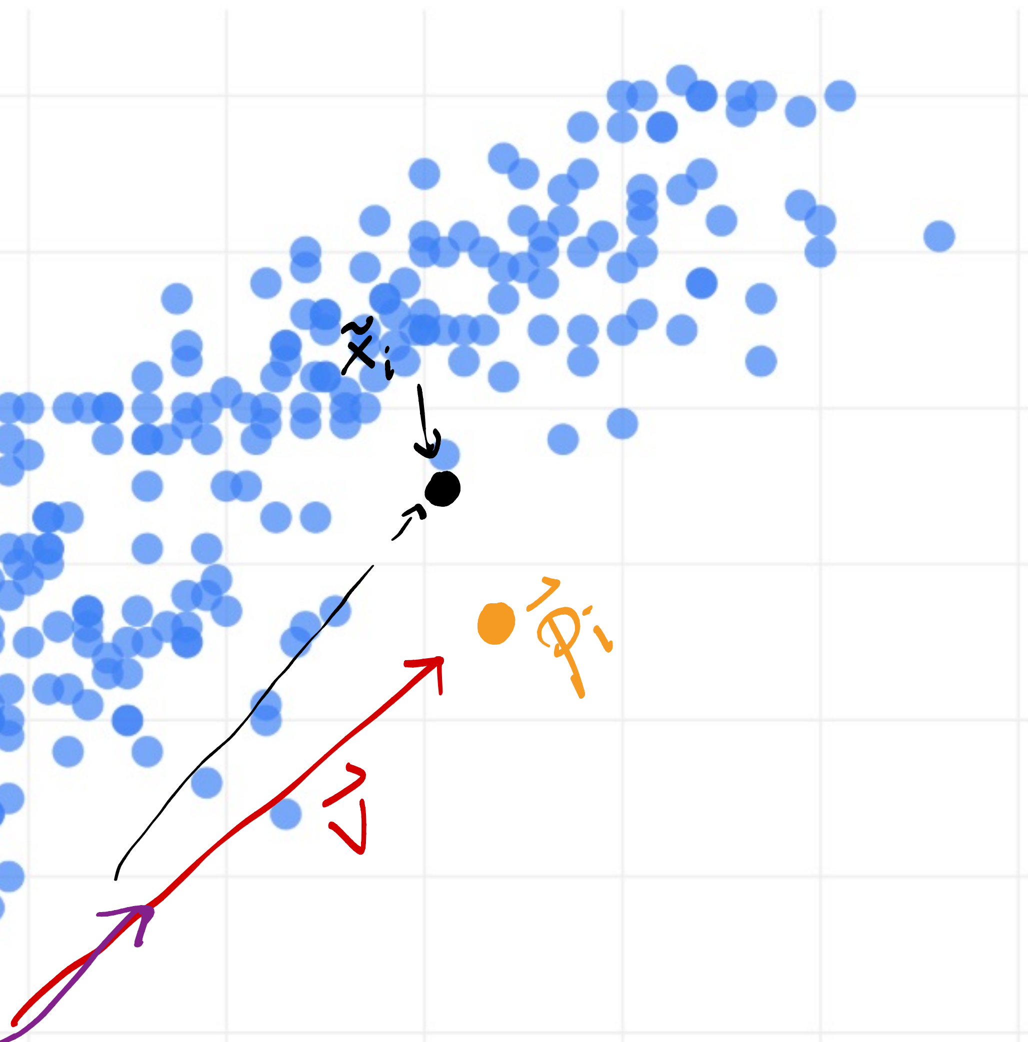


...which is why we center the data first! This doesn't change its shape.



$\tilde{X}$ = taking every column of X
and from it,
subtracting the average
of that column

"mean-centered" version
of X



(pretend this data is centered)

projection of $\tilde{x}_i$ onto $\vec{v}$

$$\vec{p}_i = \left(\frac{\tilde{x}_i \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \right) \vec{v}$$

Length of $\vec{v}$ arbitrary; assume it must be unit vector,

$$\|\vec{v}\| = 1$$

$$\Rightarrow \vec{p}_i = (\tilde{x}_i \cdot \vec{v}) \vec{v}$$

mean squared orthogonal error

$$\frac{1}{n}\sum_{i=1}^{n}\|\tilde{x}_i - \vec{p}_i\|^2$$

$$J(\vec{v}) = \frac{1}{n}\sum_{i=1}^{n}\|\tilde{x}_i - (\tilde{x}_i \cdot \vec{v})\vec{v}\|^2$$

→ best $\vec{v}$ is first column of V in $\tilde{X} = U\Sigma V^T$

row of V^T

→ aka best $\vec{v}$ = eigvec of $\tilde{X}^T\tilde{X}$ with biggest eigval!

The shorter the orthogonal errors are, the more spread out the orange points are!